

Incorporation of Fluorescent-Labeled Non-α-Amino Carboxylic Acids into the N-Terminus of Proteins in Response to Amber Initiation Codon

Masanori Miura, Norihito Muranaka, Ryoji Abe, and Takahiro Hohsaka*

School of Materials Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi 923-1292

Received December 4, 2009; E-mail: hohsaka@jaist.ac.jp

Incorporation of non-natural amino acid derivatives containing fluorescent groups into proteins is a useful method for protein analyses. Here, we investigated the incorporation of fluorescent-labeled non-α-amino carboxylic acids into the N-terminus of proteins in response to the UAG initiation codon. A series of TAMRA-labeled amino carboxylic acids were synthesized and attached to an amber suppressor initiator tRNA derived from *Escherichia coli* initiator tRNA. Fluorescent-labeled amino carboxylic acids were successfully incorporated into the N-terminus of streptavidin by adding TAMRA-acylated initiator tRNAs to an *E. coli* cell-free translation system, although the incorporation efficiency differed depending on the amino carboxylic acid chain length. Based on this observation, 3-aminopropionic acid derivatives labeled with BODIPY, rhodamine, and cyanine fluorophores were designed and synthesized. The fluorescent-labeled 3-aminopropionic acid derivatives developed using BODIPY and rhodamine dyes could be incorporated with good efficiency. On the other hand, 6-aminohexanoic acid was effectively incorporated when labeled with cyanine dyes. The present study demonstrates that translation initiation can accept a wide variety of non-natural substrates and provides a useful method for N-terminal-specific labeling of proteins with various fluorophores.

Fluorescence labeling is a very useful method for analyzing protein structure and function. Chemical modification using fluorescent molecules has generally been used for protein fluorescence labeling. However, most proteins have multiple reactive residues such as lysine and cysteine on their surfaces, and therefore site-directed fluorescence labeling is not easily achieved. Chemical modification of engineered cysteines facilitates site-directed labeling, although quantitative labeling remains difficult. Fusion gene expression with green fluorescent protein (GFP) derivatives is an alternative method for site-directed fluorescence labeling, but the large molecular size of these derivatives may influence the structure and function of fused proteins.

On the other hand, site-directed and quantitative incorporation of fluorescent groups into proteins using non-natural amino acid mutagenesis[1–4] has been developed. Various fluorescent non-natural amino acids have been incorporated in response to an amber codon and four-base codons.[5–12] This method enables the introduction of fluorescent groups at desired positions. However, incorporation into inappropriate positions may significantly damage proteins. N- or C-Terminal-specific fluorescence labeling is desirable as a general method available for labeling a wide variety of proteins without causing significant damage.

Ribosome-mediated N-terminal-specific labeling has been developed using initiator tRNAs. Initiator Met–tRNAs whose α-amino groups are acylated with non-natural groups can introduce non-natural methionine derivatives into the N-terminus of proteins using cell-free translation systems.[13] Using this method, methionine derivatives that have various fluorophores[14–16] and biotin[17] labeled at their α-amino groups have been successfully incorporated. However, the N-terminal incorporation in response to the AUG initiation codon competes with the incorporation of methionine or formylated methionine by endogenous initiator tRNA, which decreases the labeling yield. In contrast, using an amber codon UAG, originally a stop codon, as an initiation codon has helped to achieve quantitative labeling without competition with endogenous initiator tRNA.[18–20] Alternatively, a reconstituted cell-free translation system without addition of methionine has been used for N-terminal-specific incorporation of non-natural amino acids and peptide derivatives.[21] The initiation codon-mediated method can achieve rapid and quantitative N-terminal-specific labeling, independent of N-terminal amino acid sequence. These properties are advantageous compared to other N-terminal-specific labeling methods such as those using protein splicing,[22,23] leucyl/phenylalanyl–tRNA–protein transferase,[24] and sortase.[25]

In addition to α-amino acid derivatives, we reported the incorporation of BODIPYFL-labeled 3-aminopropionic and 6-aminohexanoic acids without α-amino groups[26] and *p*-aminophenylalanine derivatives double-labeled with both BODIPYFL and biotin[27] into the N-terminus of proteins in response to the UAG initiation codon. These results suggest that various non-natural molecules are accepted as substrates during translation initiation. We found that *p*-substituted phenylalanine derivatives are efficient substrates for peptide elongation.[5] Based on this substrate specificity, *p*-aminophenylalanine derivatives labeled with fluorophores,[6,7] biotin,[28] and poly(ethylene glycol) chains[29] at the *p*-amino groups have been designed and successfully incorporated into proteins. As with translation elongation, it will be worthwhile to identify substrate specificities during translation initiation in order to design non-natural molecules that can be efficiently incorporated into the N-terminus of proteins.

Published on the web April 29, 2010; doi:10.1246/bcsj.20090320

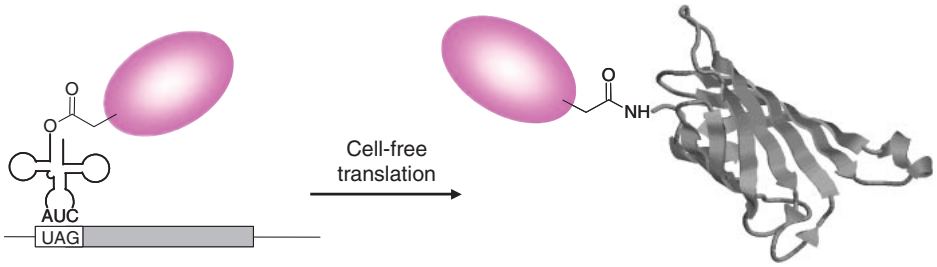

Figure 1. Schematic illustration of the incorporation of fluorescent-labeled non-α-amino carboxylic acids into the N-terminus of proteins in response to a UAG initiator codon.

In this study, we investigated the incorporation of various fluorescent-labeled non-α-amino carboxylic acids into the N-terminus of proteins in response to the UAG initiation codon in an *E. coli* cell-free translation system (Figure 1). Tetramethylrhodamine (TAMRA)-labeled amino carboxylic acids containing various alkyl chains were synthesized and their incorporations into the N-terminus of proteins were evaluated. In addition, the incorporation of amino carboxylic acid derivatives labeled with various fluorophores was shown to be useful for the N-terminal-specific fluorescence labeling of proteins.

Experimental

Materials. Succinimide esters of TAMRA, AlexaFluor488, BODIPYFL, BODIPY558, 5-carboxyrhodamine110 (CR110), and 5-carboxyfluorescein (FAM) were purchased from Molecular Probes (Eugene, OR, USA). Cy3 and Cy5 monoreactive dyes were from GE Healthcare (Piscataway, NJ, USA). T4 RNA ligase was from Takara BIO (Osaka, Japan). *E. coli* S30 extract for a linear template, alkaline phosphatase-labeled anti-mouse IgG, and MagneHis Ni-Particles were from Promega (Madison, WI, USA). RTS *E. coli* linear template generation set was from Roche Diagnostics (Basel, Switzerland). Anti-His tag antibody was from Novagen (La Jolla, CA, USA). Lysyl endopeptidase was from Wako Chemicals (Osaka, Japan). XTerra C18 and XBridge C18 columns were from Waters (Milford, MA, USA). Poros R2/10 column was from Applied Biosystems (Foster, CA, USA). ZipTip C18 was from Millipore (Milford, MA, USA).

Synthesis of Aminocarboxyl–Dinucleotides. Dioxane solution of $(Boc)_2O$ (Boc = *t*-butyloxycarbonyl) (1.5 equiv, 1.68 mmol) (200 µL) was gradually added to a mixture of amino carboxylic acid (1.12 mmol) and $NaHCO_3$ (2 equiv, 2.24 mmol) in water/dioxane (1:1 v/v, 4 mL) on ice. The mixture was stirred for 18 h on ice, acidified to pH 2 with aqueous 5% $KHSO_4$, and extracted with ethyl acetate. The extract was washed three times with aqueous 5% $KHSO_4$, once with saturated aqueous NaCl, and dried over sodium sulfate. The solvent was evaporated to give Boc–amino carboxylic acid. To a mixture of the resulting Boc–amino carboxylic acid and triethylamine (500 µL) in acetonitrile (2 mL), chloroacetonitrile (200 µL) was gradually added on ice. The mixture was stirred for 18 h on ice, acidified to pH 2 with aqueous 5% $KHSO_4$, and extracted with ethyl acetate. The extract was washed three times with aqueous 5% $KHSO_4$, three times with aqueous 4% $NaHCO_3$, once with saturated aqueous NaCl, and dried over sodium sulfate. The solvent was evaporated to give Boc–amino carboxylic acid cyanomethyl ester. Acylation of 5′-*O*-phosphoryl-2′-deoxycytidylyl(3′–5′)adenosine (pdCpA) was carried out by adding Boc–amino carboxylic acid cyanomethyl ester (8.8 µmol) to 44 mM DMF solution of pdCpA tetrabutylammonium salt (15 µL, 0.66 µmol). The resulting solution was incubated at room temperature for 1 h. Diethyl ether (600 µL) was added to the solution, and the precipitate was collected by centrifugation. The resulting precipitate was washed twice with diethyl ether and dried under vacuum. The pellet was dissolved in trifluoroacetic acid (TFA; 200 µL) and placed on ice for 15 min to remove the Boc-group. After evaporating TFA by vacuum centrifugation, the pellet was washed twice with diethyl ether (600 µL) and dried under vacuum.

Synthesis of TAMRA-Labeled Aminocarboxyl–Dinucleotides. Aqueous 0.05 M $NaHCO_3$ (60 µL) was added to a mixture of 2.2 mM dimethylsulfoxide (DMSO) solution of aminocarboxyl–pdCpA (30 µL, 66 nmol), 50 mM 5-TAMRA succinimide ester (3 µL, 150 nmol), and DMSO (27 µL). The mixture was incubated on ice for 3 h and neutralized by adding 1 M acetic acid (4.5 µL). TAMRA-labeled aminocarboxyl–pdCpA was purified using reverse-phase HPLC (XBridge C18, 2.5 µm, 4.6 × 20 mm, 1.5 mL min^{-1}, 0–100% methanol in 0.38% formic acid, over 10 min). The product was identified by electrospray ionization time-of-flight mass spectrometry (ESI-TOF MS) (Mariner, Applied Biosystems). TAMRA–AC_2–pdCpA, calculated for $(M - H)^-$ 1104.265, found 1104.322; TAMRA–AC_3–pdCpA, calculated for $(M - H)^-$ 1118.281, found 1118.236; TAMRA–AC_4–pdCpA, calculated for $(M - H)^-$ 1132.297, found 1132.209; TAMRA–AC_5–pdCpA, calculated for $(M - H)^-$ 1146.312, found 1146.326; TAMRA–AC_6–pdCpA, calculated for $(M - H)^-$ 1160.328, found 1160.322; TAMRA–AC_7–pdCpA, calculated for $(M - H)^-$ 1174.344, found 1174.290; TAMRA–AC_{12}–pdCpA, calculated for $(M - H)^-$ 1244.422, found 1244.437; TAMRA–X–AC_{12}–pdCpA; calculated for $(M - H)^-$ 1357.506, found 1357.505; TAMRA–X–Met–pdCpA, calculated for $(M - H)^-$ 1291.369, found 1291.321; TAMRA–X–AF–pdCpA, calculated for $(M - H)^-$ 1322.407, found 1322.418.

Aminopropionyl–pdCpA derivatives labeled with AlexaFluor488, BODIPYFL, BODIPY558, CR110, FAM, Cy3, and Cy5, and aminohexanoyl–pdCpA derivatives labeled with Cy3 and Cy5 were synthesized in a similar manner using the corresponding succinimide esters. ESI-TOF MS; AlexaFluor488–AC_3–pdCpA, calculated for $(M - H)^-$ 1222.133, found 1222.129; BODIPYFL–AC_3–pdCpA, calculated for $(M - H)^-$ 980.2476, found 980.2482; BODIPY558–AC_3–pdCpA, calculated for $(M - H)^-$ 1118.281, found

1118.293; FAM–AC_3–pdCpA, calculated for $(M - H)^-$ 1064.187, found 1064.202; Cy3–AC_3–pdCpA, calculated for $(M - H)^-$ 1318.335, found 1318.305; Cy5–AC_3–pdCpA, calculated for $(M - H)^-$ 1344.351, found 1344.387; Cy3–AC_6–pdCpA, calculated for $(M - H)^-$ 1361.311, found 1361.320; Cy5–AC_6–pdCpA, calculated for $(M - H)^-$ 1387.349, found 1387.284.

Preparation of Aminocarboxyl–tRNAs. An *E. coli* initiator tRNA containing a CUA anticodon and lacking the 3′-terminal dinucleotide was synthesized as described previously.[26] Ligation of the truncated initiator tRNA and fluorophore-labeled aminocarboxyl–pdCpAs was carried out in a reaction mixture (10 μL) containing 5.5 mM HEPES–Na (pH 7.5), 1 mM ATP, 15 mM $MgCl_2$, 3.3 mM dithiothreitol (DTT), 2 μg mL^{-1} bovine serum albumin (BSA), 0.25 mol tRNA, 2.2 nmol of fluorescent-labeled aminocarboxyl–pdCpA in DMSO (1 μL), and T4 RNA ligase (30 units). The reaction mixture was incubated at 4 °C for 12 h, except for TAMRA–X–AF which was incubated for 2 h in order to prevent hydrolysis of the ester bond between the amino acid and tRNA. After incubation, potassium acetate (pH 4.5) was added to a final concentration of 0.3 M. The tRNA was isolated by ethanol precipitation and dissolved in pre-chilled 1 mM potassium acetate (pH 4.5). The acylation yield was determined using reverse-phase HPLC (Poros R2/10, 4.6 × 100 mm), flow rate = 1.0 mL min^{-1} with a linear gradient of 0–40% acetonitrile in 0.1 M triethylammonium acetate (pH 7.0) over 20 min.

Cell-Free Translation. Streptavidin mRNA containing the UAG initiation codon and His-tag at N- and C-termini, respectively, was prepared as described previously.[26] The mRNA and each fluorescent-labeled aminocarboxyl–tRNA were added to an *E. coli* cell-free translation system. Translation was performed in a reaction mixture (10 μL) containing 55 mM HEPES–KOH (pH 7.5), 210 mM potassium glutamate, 6.9 mM ammonium acetate, 12 mM magnesium acetate, 1.7 mM DTT, 1.2 mM ATP, 0.28 mM GTP, 26 mM phosphor-(enol)pyruvate, 1 mM spermidine, 1.9% poly(ethylene glycol), 35 μg mL^{-1} folinic acid, 0.1 mM of 20 standard amino acids, 8 μg mRNA, 80 pmol fluorescent-labeled aminocarboxyl–tRNA, and an *E. coli* S30 extract (2 μL). The reaction mixture was incubated at 37 °C for 60 min. The reaction mixture (10 μL) was then mixed with 10 μg of RNase A (10 μL) and incubated at 37 °C for 15 min to digest the remaining acyl–tRNA. The resulting solution (2 μL) was mixed with 2× sample buffer for sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) (10 μL) and water (8 μL), and the sample (5 μL) was subjected to 20% SDS–PAGE after boiling. The fluorescent bands on the SDS–PAGE gel were visualized and quantified using a fluorescence scanner (FMBIO-III, Hitachi Software Engineering, Tokyo, Japan). The same gel was analyzed by Western blot using an anti-His tag antibody and alkaline phosphatase-labeled anti-mouse IgG. The band intensities were quantified using Scion Image program (Scion Corporation, Frederick, Maryland, USA), from which the yields of the full-length streptavidin were determined using a calibration curve prepared by measuring the band intensities of serial dilutions (100, 50, 25, 12.5, 6.25, and 3.13%) of the full-length streptavidin expressed from a wild-type mRNA.

Mass Analysis. The incorporation of TAMRA–AC_3, TAMRA–AC_4, and CR110–AC_3 was confirmed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS). The cell-free translation reaction mixture (50 μL) was loaded onto the MagneHis Ni-particles (20 μL). The beads were washed five times with 200 μL of wash buffer (10 mM Tris-HCl, pH 7.5, 100 mM NaCl), and the His-tagged protein was eluted with 25 μL of elution buffer (10 mM Tris-HCl, pH 7.5, 500 mM imidazole). The eluent was mixed with 0.01 units of lysyl endopeptidase and incubated at 30 °C for 6 h. The resulting peptide fragments were desalted and concentrated using ZipTip C18 and eluted with a matrix solution containing saturated α-cyano-4-hydroxycinnamic acid in 1:1 mixture of acetonitrile and 0.1% TFA. Mass measurement was performed by MALDI-TOF MS (Voyager DE-Pro, Applied Biosystems) in the positive mode using angiotensin II as an external calibrant.

Results and Discussion

Synthesis of TAMRA-Labeled Aminocarboxyl–tRNAs. To investigate the effects of linker structures on the incorporation of fluorescent-labeled amino carboxylic acids, a variety of TAMRA-labeled amino carboxylic acids were designed as shown in Figure 2. TAMRA was chosen as a model fluorophore because of good characteristic properties, such as high fluorescence quantum yield, long emission wavelength, and high photostability. TAMRA–AC_n, where n is 2–8 and 12, contained alkyl linkers of n carbon atoms. TAMRA–X–AC_{12} contained an aminohexyl linker in addition to an AC_{12} linker. TAMRA–EG_8 containing an ethylene oxide linker of 8 atoms was also used to examine the influence of linker flexibility. In addition, methionine (TAMRA–Met and TAMRA–X–Met) and *p*-aminophenylalanine derivatives (TAMRA–X–AF) labeled at the α- and *p*-amino groups, respectively, were also examined.

TAMRA-labeled amino carboxylic acids were attached to a dinucleotide (pdCpA) to prepare acyl–tRNAs by a chemical aminoacylation method.[30,31] For this purpose, aminocarboxyl–pdCpA derivatives were first synthesized via cyanomethyl esters after which they were coupled with TAMRA succinimide ester. This synthetic route was advantageous for increasing the reaction yield using expensive fluorophore succinimide esters and synthesizing aminocarboxyl–pdCpAs labeled with various fluorophores. All products were purified by HPLC, identified by ESI-TOF MS, and ligated with a synthetic *E. coli* initiator tRNA containing the CUA anticodon.[26] HPLC analysis of the ligated tRNAs using a Poros R2/10 column[16] showed that the ligation reaction efficiencies, independent of the amino carboxylic acid types used, were about 80%.

Incorporation of TAMRA-Labeled Amino Carboxylic Acids into the N-Terminus of Streptavidin. The incorporation of TAMRA-labeled amino carboxylic acids into proteins was examined using a streptavidin gene containing the UAG initiation codon and His-tag at N- and C-termini, respectively. When the UAG initiation codon is decoded by the initiator tRNA acylated with TAMRA-labeled amino carboxylic acids, TAMRA-containing streptavidin will be obtained. On the other hand, when TAMRA-labeled amino carboxylic acids are not incorporated into the protein, no proteins will be produced because the translation initiation does not occur. Therefore, the

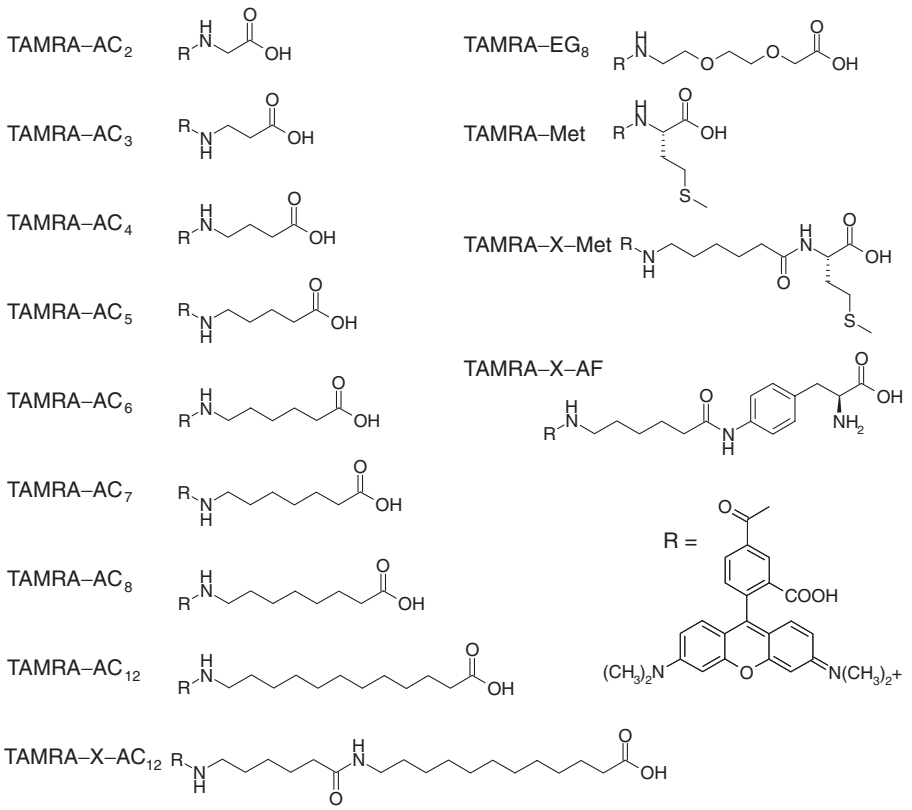

Figure 2. Structures of TAMRA-labeled non-α-amino carboxylic acids and α-amino acids.

incorporation of TAMRA-labeled amino carboxylic acids can be evaluated by judging the expression level of the TAMRA-containing streptavidin.

The streptavidin mRNA was expressed in an *E. coli* cell-free translation system in the presence of TAMRA-labeled aminocarboxyl–tRNAs. Cell-free translation products were separated by SDS–PAGE and visualized by fluorescence imaging, and the same gel was analyzed by Western blot using anti-His tag antibody. The fluorescence image showed that fluorescent bands around 17 kDa, which corresponded to monomeric streptavidin, were observed for all TAMRA-labeled amino carboxylic acids (Figure 3a), though TAMRA–AC_2 gave a very faint band. The Western blot also showed that full-length translation products were obtained in the presence of TAMRA-labeled aminocarboxyl–tRNAs. On the other hand, no bands were observed in the absence of the amber suppressor initiator tRNA, indicating that the UAG initiation codon was not decoded by any tRNA other than the added initiator tRNAs. These results indicated that the streptavidins obtained in the presence of TAMRA-labeled aminocarboxyl–tRNAs were quantitatively labeled with TAMRA at the N-terminus.

The translation products obtained in the presence of TAMRA–AC_3–tRNA and TAMRA–AC_4–tRNA were analyzed by MALDI-TOF MS after purification on Ni–NTA beads and digestion with lysyl endopeptidase. As expected, a mass peak corresponding to TAMRA–AC_3-containing N-terminal peptides (TAMRA–AC_3–Asp–Pro–Ser–Lys; calculated for MH^+ 929.4, found 929.5) was observed (Figure 4a). For TAMRA–AC_4, a mass value higher by 14 was observed (calculated for MH^+ 943.4, found 943.5; Figure 4b), which corresponded to the increase in the alkyl chain length. These results supported that the TAMRA-labeled amino carboxylic acids were evidently incorporated into the N-terminus of the protein. The mass spectra also showed weak peaks with mass values lower by 14 and 44 than the TAMRA-containing peptides, which could be identified as TAMRA-containing peptides lacking a methyl group and a carboxyl group of TAMRA, respectively, because TAMRA succinimide ester also gave weak peaks with mass values lower by 14 and 44 possibly due to photodegradation of TAMRA during the laser-ionization in the MALDI-TOF MS measurement.

Fluorescence band intensities were quantified to compare the incorporation efficiencies of TAMRA-labeled amino carboxylic acids (Figure 3b), assuming that the fluorescence intensity was not affected by the type of amino carboxylic acids. The incorporation efficiencies differed depending on the structure and chain length of amino carboxylic acids. From TAMRA–AC_3 to AC_7, the incorporation efficiency decreased with increasing chain length. AC_{12} and X–AC_{12} derivatives also showed decreasing efficiencies with increasing chain length. However, TAMRA–AC_8 showed an inordinately high efficiency. EG_8 derivative containing the same 8-atom chain as AC_8 was incorporated with an efficiency comparable to AC_8, suggesting that linker flexibility was not a significant factor. TAMRA–Met and X–Met derivatives showed higher incorporation efficiencies than most of the amino carboxylic acid derivatives.

The relationship between the structure of amino carboxylic acids and incorporation efficiency suggests the presence of a binding space for TAMRA near the tRNA, for derivatives with

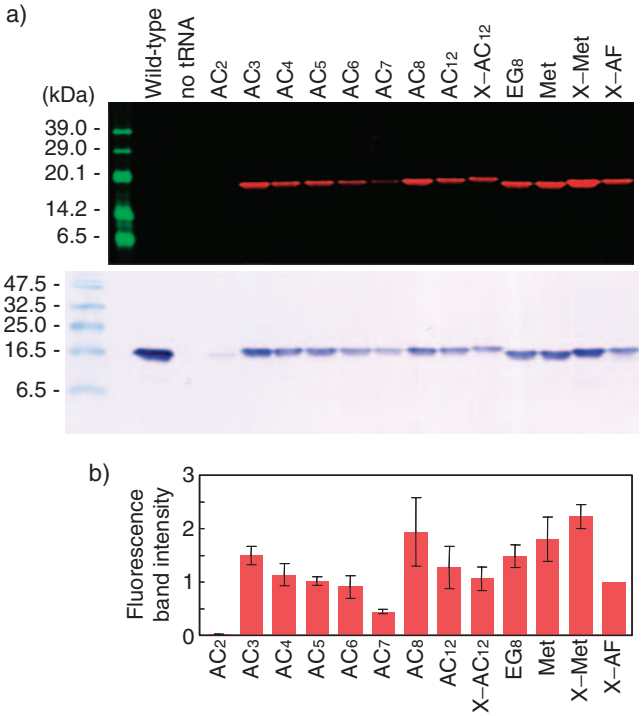

Figure 3. Incorporation of TAMRA-labeled amino carboxylic acids into the N-terminus of streptavidin in response to the UAG initiation codon. (a) Fluorescence image of 20% SDS–PAGE gel with excitation at 532 nm and emission at 580 nm, and Western blot analysis of the same gel using anti-His tag antibody. (b) Fluorescence band intensities for TAMRA-labeled streptavidin observed in the SDS–PAGE gel fluorescence image. Results are mean ± SDs of three assays.

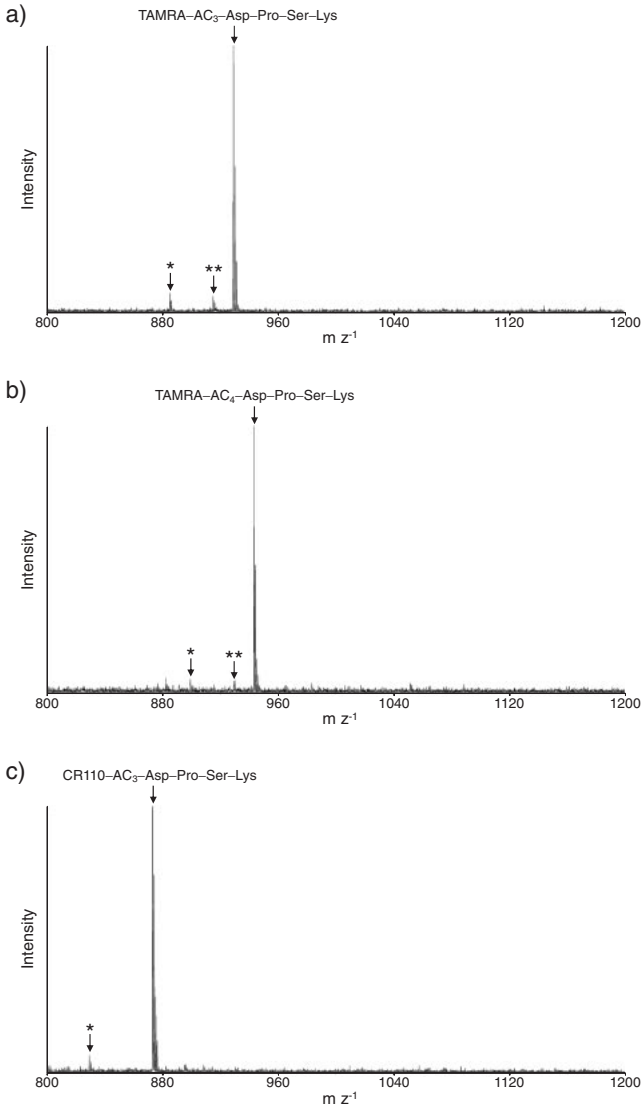

Figure 4. MALDI-TOF MS measurements of the lysyl endopeptidase-digested N-terminal peptide fragments (Asp–Pro–Ser–Lys) labeled with TAMRA–AC_3 (calculated for MH^+ 929.4, observed 929.6), TAMRA–AC_4 (calculated 943.4 for MH^+, found 943.5), and CR110–AC_3 (calculated 873.3 for MH^+, found 873.5). Peaks indicated by asterisks (∗ and ∗∗) could be identified as labeled peptide fragments lacking a carboxyl group and a methyl group, respectively, which were generated possibly due to photodegradation of the fluorophores during the laser-ionization.

less than 8-carbon chains. In TAMRA–AC_8, EG_8, and much longer derivatives, TAMRA may be located at another binding space far from the tRNA or may be surface-exposed. The incorporation of TAMRA–X–AC_{12} with a 19-atom chain revealed that very large non-natural molecules are potentially acceptable substrates during ribosomal initiation.

The highly efficient incorporation of TAMRA–Met and X–Met may have been due to the binding of the methionine side chain to the binding site originally meant for formylmethionine. Nonetheless, TAMRA–AC_3 and AC_8 showed nearly comparable efficiencies for the methionine derivatives with similar length linkers, suggesting that fluorescent-labeled amino carboxylic acids can be efficiently incorporated without contributions from the binding of the methionine side chain and α-amino or amide groups.

The inefficient incorporation of TAMRA–AC_2 may be due to the short alkyl groups without side-chains being inappropriate for TAMRA to be accepted by the translation machinery. In addition, the α-amide bond may have destabilized the aminoacyl linkage between TAMRA–AC_2 and tRNA as suggested in previous study,[26] and enhanced the hydrolysis of this linkage. To evaluate the stability of the ester bond between the TAMRA-labeled amino carboxylic acids and the initiator tRNA, the hydrolysis of TAMRA–AC_2– and TAMRA–AC_3–tRNAs in the same pH and temperature conditions (pH 7.5 at 37 °C) as for the cell-free translation was monitored using reverse-phase HPLC (Figure 5). The hydrolysis yields of TAMRA–AC_2–tRNA were 10% at 1 h and 95% at 24 h, while TAMRA–AC_3–tRNA showed negligible hydrolysis at 1 h and 36% hydrolysis at 24 h. These results showed that the α-amide bond evidently destabilized the ester bond. However, the result that 90% of TAMRA–AC_2–tRNA remained after 1 h was not enough to explain the inefficient incorporation of TAMRA–AC_2. Therefore, the destabilized ester bond may only partially contribute to the inefficient incorporation of TAMRA–AC_2.

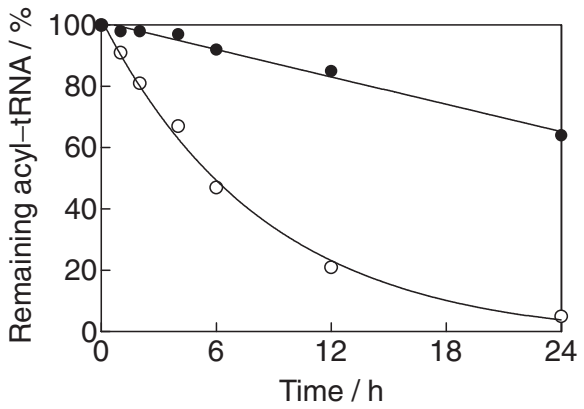

Figure 5. Time course of the hydrolysis of TAMRA–AC_2–tRNA (open circle) and TAMRA–AC_3–tRNA (filled circle) in 10 mM Tris-HCl (pH 7.5) at 37 °C. The remaining TAMRA–tRNAs were monitored by reverse-phase HPLC.

The concentration dependence of incorporation efficiencies of TAMRA-labeled aminocarboxyl–tRNAs was examined. Nearly the same amounts of TAMRA-labeled streptavidin were obtained for all derivatives in the presence of 40, 80, and 160 pmol of TAMRA-labeled aminocarboxyl–tRNAs in 10 μL of cell-free translation, indicating that tRNA concentration (80 pmol per 10 μL of cell-free translation) was sufficient.

Incorporation of 3-Aminopropionic Acids with Various Fluorophores. We reported that BODIPYFL-labeled 3-aminopropionic acid is efficiently incorporated into the N-terminus of proteins. The present study also showed that 3-aminopropionic acid is an effective framework for incorporating TAMRA. These results suggested that the aminopropionic acid framework would be effective for introducing various fluorophores. To demonstrate this, 3-aminopropionic acid derivatives labeled with various fluorophores, such as AlexaFluor488, 5-carboxyrhodamine110 (CR110), 5-carboxyfluorescein (FAM), BODIPY dyes, and Cy dyes (Figure 6a), were synthesized and their incorporations into the N-terminus of streptavidin were examined.

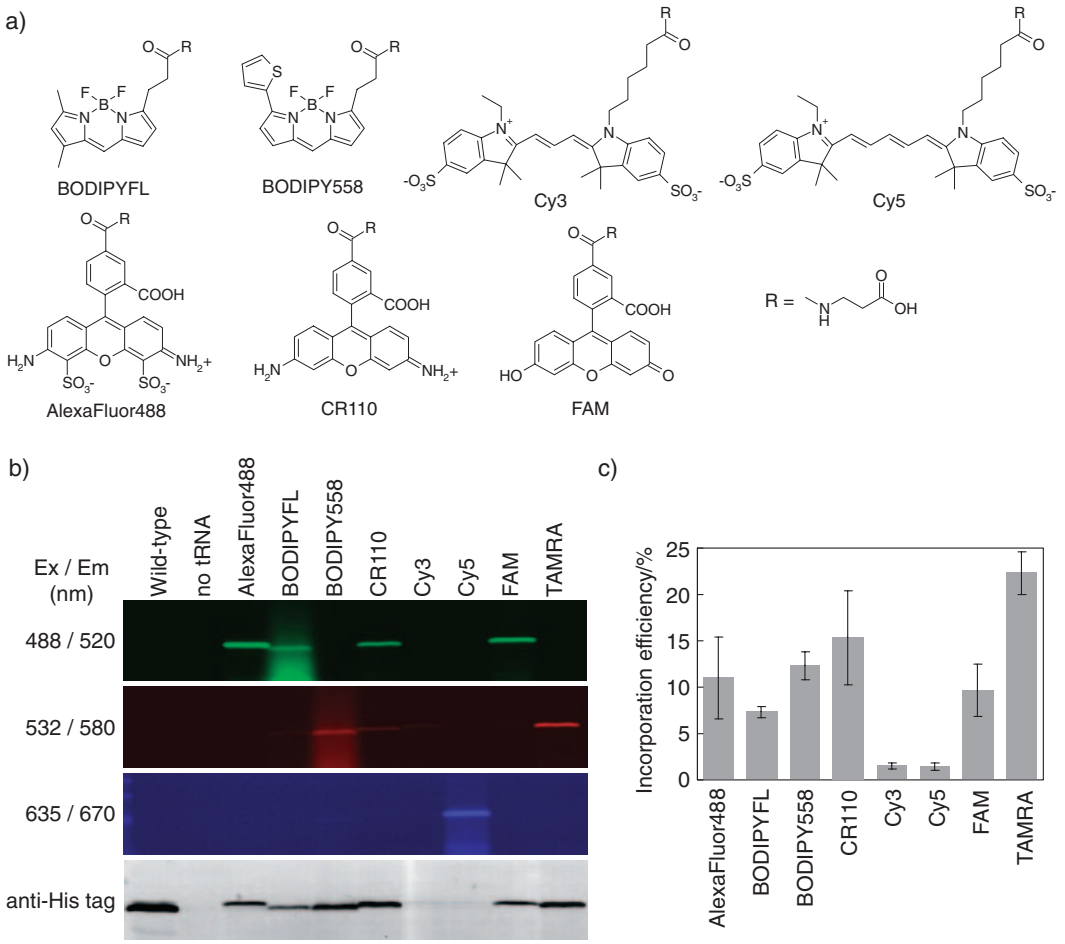

Figure 6. Incorporation of various fluorescent-labeled 3-aminopropionic acids into the N-terminus of streptavidin in response to the UAG initiation codon. (a) Structures of fluorescent-labeled 3-aminopropionic acids. (b) Fluorescence image of 20% SDS–PAGE gel with excitation and emission at 488 and 520 nm (green), 532 and 580 nm (red), and 635 and 670 nm (blue), and Western blot of the same SDS–PAGE gel using anti-His tag antibody. (c) Incorporation efficiencies of the fluorescent-labeled 3-aminopropionic acids into the N-terminus of streptavidin determined by comparing band intensities on Western blot with those of serial dilutions of wild-type streptavidin. Results are mean ± SDs of three assays.

Fluorescence images of SDS–PAGE showed that fluorescent-labeled 3-aminopropionic acids, other than Cy3 derivative, were successfully incorporated into the N-terminus of streptavidin (Figure 6b). It should be noted that smear bands were observed for BODIPYFL and BODIPY558 at the lower molecular-weight range on the SDS–PAGE, which were derived from low-molecular-weight BODIPY-labeled aminopropionic acids due to the tendency of the hydrophobic BODIPY dyes to adsorb to the polyacrylamide gel. For AlexaFluor488 and FAM, weak bands were observed at the lower molecular-weight range, which may be generated by any irregular termination in the peptide elongation process. A similar truncated polypeptide had been observed on Western blotting using an anti-T7 tag antibody for wild-type streptavidin containing T7 tag at the N-terminus.[5] The truncated by-products generated in the peptide elongation process would not affect the incorporation of the aminocarboxylic acids in the initiation process and, in addition, could be removed by purification using Ni chelate affinity chromatography. However, the generation of the truncated by-products slightly decreased the yields of the full-length proteins. From the fluorescence band intensities, the relative yields of the truncated polypeptides were estimated to be 5 and 8% compared to the full-length proteins for AlexaFluor488 and FAM, respectively. Optimization of the nucleotide sequence may be effective to suppress the irregular termination of the peptide elongation.

MALDI-TOF MS analysis of the lysyl endopeptidase-digested translation product for CR110–AC_3 showed that the expected mass peak was evidently observed (calculated for MH^+ 873.3, found 873.5; Figure 4c) as in the case of TAMRA. The mass spectrum also showed a weak peak with mass value lower by 44, which could be identified as a CR110-containing peptide lacking a carboxyl group of CR110 due to photodegradation during the laser-ionization. Unlike the case of TAMRA, CR110, which contained no methyl group in the fluorophore, gave no peak with mass value lower by 14, supporting the idea that the mass value lower by 14 observed for TAMRA is due to the lack of the methyl group.

The relative yields of the proteins were quantified by comparing the band intensities on Western blot with those of serial dilutions of wild-type (Figure 6c). Fluorescein and rhodamine derivatives showed about 10–20% yields, and TAMRA derivative showed the highest yield (22%). Cy5 derivative showed very low yield, whereas a distinct fluorescent band was observed, possibly due to its bright fluorescence. These results indicate that various fluorescent groups are accepted as substrates during translation initiation, and the incorporation efficiency depended on the structure of the fluorescent groups. In contrast, Cy3 and Cy5 derivatives were incorporated with very low efficiencies, suggesting that the large molecular size of Cy dyes may cause steric hindrance with the ribosomal initiation complex.

Influence of Chain Length on Incorporation of Cy Dyes. In the case of TAMRA, an 8-aminooctanoic acid derivative (TAMRA–AC_8) was incorporated more efficiently than amino carboxylic acids with shorter alkyl chains. This raised the possibility that Cy dyes, which originally have a hexyl linker, may be incorporated by increasing the amino-carboxyl chain length. To test this possibility, Cy3- and Cy5-labeled 6-aminohexanoic acid derivatives (Cy3– and Cy5–AC_6) were synthesized and their incorporations were investigated.

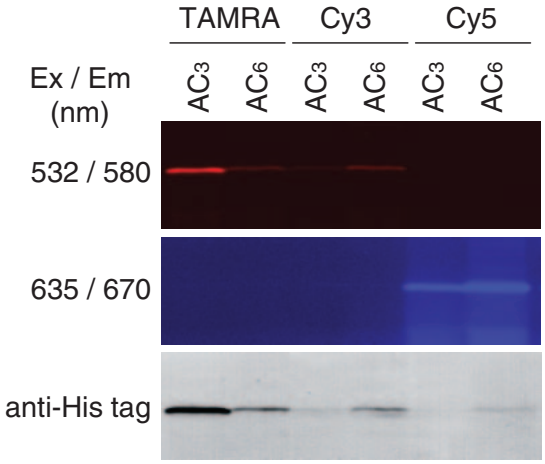

Figure 7. Incorporation of TAMRA-, Cy3-, and Cy5-labeled 3-aminopropionic acid (AC_3) and 6-aminohexanoic acid (AC_6) derivatives into the N-terminus of streptavidin in response the UAG initiation codon. Fluorescence image of 20% SDS–PAGE gel with excitation and emission at 532 and 580 nm (red) and 635 and 670 nm (blue), and Western blot of the same SDS–PAGE gel using anti-His tag antibody.

A fluorescence image of SDS–PAGE and Western blotting showed that the incorporations of Cy3 and Cy5 were enhanced by attaching them to the aminohexanoic acid (Figure 7), and the incorporation efficiencies estimated by Western blot were about 3%. The aminohexyl chain may be effective in delivering Cy dyes to the above-mentioned binding space which was far from the acyl–tRNAs or to surface-exposed sites. This finding will be valuable for designing amino carboxylic acids containing large fluorophores.

Incorporation into Other Proteins. The incorporation of TAMRA–AC_3 was examined for proteins other than streptavidin. As shown in Figure 8, TAMRA–AC_3 was successfully incorporated into maltose-binding protein (MBP) and interleukin 2 (IL2) in response to the UAG initiator codon, suggesting that fluorescent-labeled non-α-amino acids can be incorporated into the N-terminus of proteins regardless of N-terminal amino acid sequence. A weak band observed at the lower molecular-weight range may be generated by any irregular termination in the peptide elongation process as in the case of the incorporation of AlexaFluor488 and FAM into streptavidin.

Conclusion

The present study demonstrates that fluorescent-labeled non-α-amino carboxylic acids are successfully incorporated into the N-terminus of proteins in response to the UAG initiation codon. This indicates that translation initiation can accept a wide variety of non-natural substrates. The results that incorporation efficiency depends on the linker length of amino carboxylic acids and structure of fluorophores suggest that

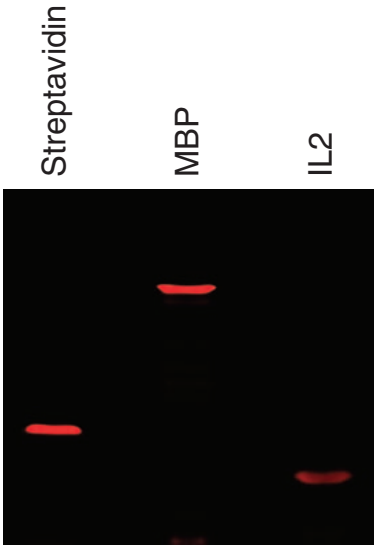

Figure 8. Fluorescence image of SDS–PAGE gel for the incorporation of TAMRA–AC_3 into the N-terminus of streptavidin, maltose-binding protein (MBP), and interleukin 2 (IL2) with excitation at 532 nm and emission at 605 nm.

various non-natural molecules can be incorporated into the N-terminus using appropriately designed structure of non-α-amino carboxylic acids. Relatively long chains would be required for non-natural molecules that originally have long alkyl chains, as in the case of Cy dyes.

This study also demonstrates that the initiation codon-mediated method allows rapid, quantitative, and neighboring sequence-independent labeling. These advantages are not achieved by other chemical or enzymatic methods including those using protein splicing, leucyl/phenylalanyl–tRNA–protein transferase, and sortase, which require particular amino acid residues at the N-terminus and are unable to achieve fully quantitative labeling. The present method will be useful for double-fluorescent-labeling and FRET analysis of proteins by combining four-base codon-mediated incorporation of fluorescent-labeled non-natural α-amino acids which allows quantitative labeling at specific internal positions.[8]

On the other hand, low-productivity of the cell-free translation system is a disadvantage of the present method that must be overcome for future applications. Nonetheless, the use of bright fluorophores can partially compensate for the low-yield of the labeled proteins. In addition, the result that non-α-aminoacyl–tRNAs are more stable against hydrolysis than α-aminoacyl–tRNAs will be advantageous for increasing the yield of fluorescent-labeled proteins in an *E. coli* dialysis cell-free translation system or wheat-germ system, which require long incubation periods.

We thank Dr. Kenzo Fujimoto for MALDI-TOF MS measurements. This work was supported by Grants-in-Aid for Scientific Research (B) (No. 19370063) and Scientific Research on Innovative Areas (No. 20107005) from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Reference

1 C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith, P. G. Schultz, *Science* **1989**, *244*, 182.
2 J. D. Bain, E. S. Diala, C. G. Glabe, T. A. Dix, A. R. Chemberlin, *J. Am. Chem. Soc.* **1989**, *111*, 8013.
3 T. Hohsaka, Y. Ashizuka, H. Murakami, M. Sisido, *J. Am. Chem. Soc.* **1996**, *118*, 9778.
4 T. Hohsaka, *Bull. Chem. Soc. Jpn.* **2004**, *77*, 1041.
5 T. Hohsaka, D. Kajihara, Y. Ashizuka, H. Murakami, M. Sisido, *J. Am. Chem. Soc.* **1999**, *121*, 34.
6 H. Murakami, T. Hohsaka, Y. Ashizuka, K. Hashimoto, M. Sisido, *Biomacromolecules* **2000**, *1*, 118.
7 T. Hohsaka, N. Muranaka, C. Komiyama, K. Matsui, S. Takaura, R. Abe, H. Murakami, M. Sisido, *FEBS Lett.* **2004**, *560*, 173.
8 D. Kajihara, R. Abe, I. Iijima, C. Komiyama, M. Sisido, T. Hohsaka, *Nat. Methods* **2006**, *3*, 923.
9 V. W. Cornish, D. R. Benson, C. A. Altenbach, K. Hideg, W. L. Hubbell, P. G. Schultz, *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 2910.
10 G. Turcatti, K. Nemeth, M. D. Edgerton, U. Meseth, F. Talabot, M. Peitsch, J. Knowles, H. Vogel, A. Chollet, *J. Biol. Chem.* **1996**, *271*, 19991.
11 B. E. Cohen, T. B. McAnaney, E. S. Park, Y. N. Jan, S. G. Boxer, L. Y. Jan, *Science* **2002**, *296*, 1700.
12 R. D. Anderson, III, J. Zhou, S. M. Hecht, *J. Am. Chem. Soc.* **2002**, *124*, 9674.
13 W. Kudlicki, O. W. Odom, G. Kramer, B. Hardesty, *J. Mol. Biol.* **1994**, *244*, 319.
14 V. Ramachandiran, C. Willms, G. Kramer, B. Hardesty, *J. Biol. Chem.* **2000**, *275*, 1781.
15 B. McIntosh, V. Ramachandiran, G. Kramer, B. Hardesty, *Biochimie* **2000**, *82*, 167.
16 S. Gite, S. Mamaev, J. Olejnik, K. Rothschild, *Anal. Biochem.* **2000**, *279*, 218.
17 M. Taki, S. Y. Sawata, K. Taira, *J. Biosci. Bioeng.* **2001**, *92*, 149.
18 U. Varshney, U. L. RajBhandary, *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 1586.
19 S. Mamaev, J. Olejnik, E. K. Olejnik, K. J. Rothschild, *Anal. Biochem.* **2004**, *326*, 25.
20 J. Olejnik, S. Gite, S. Mamaev, K. J. Rothschild, *Methods* **2005**, *36*, 252.
21 Y. Goto, H. Suga, *J. Am. Chem. Soc.* **2009**, *131*, 5040.
22 B. Schuler, L. K. Pannell, *Bioconjugate Chem.* **2002**, *13*, 1039.
23 V. Muralidharan, T. W. Muir, *Nat. Methods* **2006**, *3*, 429.
24 M. Taki, H. Kuroiwa, M. Sisido, *ChemBioChem* **2008**, *9*, 719.
25 T. Yamamoto, T. Nagamune, *Chem. Commun.* **2009**, 1022.
26 N. Muranaka, M. Miura, H. Taira, T. Hohsaka, *ChemBioChem* **2007**, *8*, 1650.
27 T. Egusa, K. Shiraga, T. Chiba, Y. Tokuda, I. Iijima, T. Hohsaka, *Bull. Chem. Soc. Jpn.* **2010**, *83*, 176.
28 T. Watanabe, N. Muranaka, I. Iijima, T. Hohsaka, *Biochem. Biophys. Res. Commun.* **2007**, *361*, 794.
29 N. Shozen, I. Iijima, T. Hohsaka, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4909.
30 S. M. Hecht, B. L. Alford, Y. Kuroda, S. Kitano, *J. Biol. Chem.* **1978**, *253*, 4517.
31 S. A. Robertson, J. A. Ellman, P. G. Schultz, *J. Am. Chem. Soc.* **1991**, *113*, 2722.